DOCUMENT RESUME

ED 420 677 TM 028 352

AUTHOR Nitko, Anthony J.; Amedahe, Francis; Al-Sarimi, Abdullah; Wang, Shudong; Wingert, Melissa

TITLE How Well Are the Kentucky Academic Expectations Matched to the KIRIS96 Assessments, CTBS, and CAT?

PUB DATE 1998-00-00

NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Academic Achievement; Achievement Tests; *Educational Objectives; Elementary Secondary Education; State Programs; Student Evaluation; *Test Items; Test Use; *Testing Programs; Thinking Skills

IDENTIFIERS California Achievement Tests; Comprehensive Tests of Basic Skills; Kentucky; *Kentucky Instructional Results Information System

ABSTRACT

Matches between questions in the Kentucky Instructional Results Information System for 1995-96 (KIRIS96), the Comprehensive Tests of Basic Skills, Fourth Edition (CTBS4), and the California Achievement Tests, Fifth Edition (CAT5) and the Kentucky Academic Expectations as they are defined in the Kentucky curriculum framework were studied. Each assessment question was read by two or three researchers familiar with the Academic Expectations. They found the KIRIS96 questions to be well written, and that they covered the Academic Expectations for the most part. Parts of the curriculum that were not so well sampled by the KIRIS96 questions were identified. The CTBS4 and the CAT5 did not cover the Kentucky curriculum framework so well. The types of questions asked by those tests required less integration of complex thinking processes than the KIRIS96 questions. However, researchers noted that the CTBS4 and the CAT5 did a better job than the KIRIS96 in separately assessing student's learning of certain traditional basic skills in spelling, mathematics computation, recognition of mathematics concepts, language mechanics, and use of specific language expressions. In general, standardized assessments have the advantage in providing measures of individual students' strengths and weaknesses in several traditional areas, and they have the advantage of being referenced to a national sample of students. (Contains three tables.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 420 677

Revised 4/7/98

How Well Are the Kentucky Academic Expectations Matched to the KIRIS96 Assessments, CTBS, and CAT? [1,2,3]

Anthony J. Nitko, Francis Amedahe, Abdullah Al-Sarimi, Shudong Wang, and Melissa Wingert
School of Education, University of Pittsburgh

Summary

In this research we studied the questions in the KIRIS for 1995-1996 (KIRIS96), the *Comprehensive Tests of Basic Skills*, Fourth Edition (CTBS4), and the *California Achievement Tests*, Fifth Edition (CAT5) in relation to the Kentucky Academic Expectations as these are defined in the Kentucky curriculum framework. We found that the KIRIS96 questions were very well written and for the most part covered the Academic Expectations. We identified parts of the curriculum framework that were not so well sampled by the KIRIS96 questions. We also found that the CTBS4 and the CAT5 did not cover the Kentucky curriculum framework well and that the types of questions asked by these tests require students to use less integration of complex thinking processes than the KIRIS96 questions. However, we noted that the CTBS4 and the CAT5 do a better job than the KIRIS96 in separately assessing student's learning of certain traditional basic skills in spelling, mathematics computation, recognition of mathematics concepts, language mechanics, and use of specific language expressions.

Standardized achievement tests, in general, have the advantage over the KIRIS96 cognitive assessments in providing measures of individual students' strengths and weaknesses in several traditional curriculum areas. They also provide measures of individual students' long term educational growth and development with respect to these traditional curriculum areas. Third, the scores from these tests are referenced to a national sample of students.

The Focus of This Study

We studied how well questions on three assessment instruments matched the Academic Expectations stated in Goals 1, 2, 5, and 6 of the Commonwealth of Kentucky's curriculum framework. This framework is described in Volume I of the document, *Transformations: Kentucky's Curriculum Framework* (September 1995). This document was developed by the Kentucky Department of Education in cooperation with panels of teachers, curriculum experts, parents, and other citizens.

The six goals described in *Transformations* are listed below. The subgoals that constitute the Academic Expectations are shown in Tables 1 through 3.

Goal 1. Use Basic Communication and Mathematics Skills

---

[1] A paper presented at the Annual Meeting of the American Educational Research Association, San Diego, April 1998.

[2] This study was conducted under a contract with the Kentucky Institute for Educational Research.

[3] We thank Brian Gong, Suzanne Lane, and Roger Pankratz for insightful comments on an earlier version of the manuscript. We are responsible for any remaining errors and misinterpretations, however.

TM028352

Anthony Nitko

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Goal 2.	Apply Core Concepts and Principles
Goal 3.	Become Self-Sufficient
Goal 4.	Become Responsible Group Members
Goal 5.	Think and Solve Problems
Goal 6.	Connect and Integrate Knowledge

Goals 3 and 4 are not part of the KIRIS96 assessment program. They were eliminated by the 1992 General Assembly following a recommendation by the then Commissioner of Education, Thomas Boyson.

The Assessments Studied

The three instruments studied were the KIRIS paper-and-pencil open-response assessments for 1995 - 1996 (KIRIS96), the *Comprehensive Tests of Basic Skills*, Fourth Edition (CTBS4), and the *California Achievement Tests*, Fifth Edition (CAT5).

KIRIS96

For the KIRIS96 assessment, we studied the Reading, Mathematics, Science, Social Studies, Arts and Humanities, and Practical Living/Vocational Studies questions. We did not study the Writing Portfolio or the Performance Events.

We studied all twelve of the forms administered to fourth, eighth, and eleventh grade children. Some questions are common (identical) across all twelve forms. Each form also contains unique questions: That is, questions that differ from one form to the next. At the fourth grade level, 164 questions were studied; at the eighth grade level 168 were studied; and eleventh grade level another 168 questions were studied.

An individual student is administered only one form, containing from 32 to 36 different questions. Different students take different forms so that within a school, all twelve forms are administered at each grade level. This plan of assessment whereby different students take different questions is commonly used in assessment when the purpose of the assessment is to evaluate groups rather than evaluate individual students. It is called *item sampling* or *matrix sampling*.

We treated all the KIRIS96 questions at a grade level as belonging to a single pool. We report our results based on this pool, rather than reporting on each form separately. The purpose of the KIRIS96 cognitive assessment is to assess Goals 1, 2, 5, and 6 as completely as possible for a school rather than to assess individual students. Thus, our results apply to how well the questions assess the school rather than how well the questions assess an individual student.

CTBS4

For the CTBS4, we studied the Complete Battery of Form A at Level 14 (fourth grade), Level 17/18 (eighth grade), and Level 21/22 (eleventh grade). The CTBS4 is a multiple-choice, standardized test that is used in many Kentucky school districts. Unlike the KIRIS96, this test focuses on traditional curriculum content. We studied all of the subtests: Vocabulary, (reading) Comprehension, Spelling, Language Mechanics, Language Expression, Mathematics Computation, Mathematics Concepts and Applications, Study Skills, Science, and Social Studies. No open-response questions appear on the CTBS4. There were 408 test questions studied at each of the three grade levels. The purpose of the CTBS4 is to assess each student in all areas, so each student who is administered the Complete Battery takes all 408 questions over several days.

CAT5

For the CAT5, we studied the Survey Battery of Form A at Level 14 (fourth grade), Level 18 (eighth grade), and Level 21/22 (eleventh grade). Like the CTBS4, the CAT5 is a multiple-choice, standardized achievement test. It is used in many Kentucky school systems as well. It also focuses on traditional curriculum content. We studied all of the subtests: Vocabulary, (reading) Comprehension, Spelling, Language Mechanics, Language Expression, Mathematics Computation, Mathematics Concepts and Applications, Study Skills, Science, and Social Studies. No open-response questions appear on the CAT5. There were 200 test questions studied at each of the three grade levels. The purpose of the CAT5 is to assess each student in all areas so each student who is administered the Survey Battery takes all 200 questions over several days.

How We Went About Studying the Assessment Questions

After we became familiar with the Academic Expectations, we read each assessment question and classified it as assessing one or more of the Academic Expectations. Each question was classified independently by two or three persons. When disagreements occurred, the question was discussed until a resolution was reached as to its classification. Many questions were classified into more than one category because they were judged to assess more than one Academic Expectation. To classify a question we relied on the detailed definitions of the Academic Expectations found

within the text of Volume I of the above-mentioned Kentucky curriculum framework document.

What We Found

Our findings are summarized in Tables 1, 2, and 3. These tables show the percentage of questions from the KIRIS96, the CTBS4, and the CAT5 questions that we judged to match each Academic Expectation. The percentages within a column do not sum to 100% because questions matched more than one Academic Expectation and were so classified. Each percentage in the table is calculated by dividing the number of questions classified into a category by the total number of distinct questions included in its respective assessment. For example, Table 1 shows that of the 164 fourth grade level KIRIS96 questions, 4% assessed Academic Expectation 2.19 Geography. If none of the questions matched a particular Academic Expectation, the cell in the table was left blank.

Conclusions

The main conclusions we draw from our study of the assessment questions are given below.

- After examining the KIRIS96 questions we note that they are very well designed and that they present students with very challenging tasks. They often present students with realistic and applied situations to which they are to respond. Reading passages, for example, are often quite long and taken from original sources. They may be termed "authentic" or "realistic" reading materials. The questions in all areas require students to use complex thinking and several intellectual skills in an integrated way in order to respond well.

- The KIRIS96 open-response questions, when viewed as a complete set, do assess most of the Academic Expectations. There are some notable exceptions, however.

- Under Goal 1, as we would expect, *visual arts*, *music*, *movement*, and *technology* are not assessed by any of the three paper-and-pencil tests. Another Academic Expectation under Goal 1, *accessing information*, is not assessed well either, at least in the way which that ability is defined in the curriculum framework. None of the KIRIS96 questions, and only a few of the CTBS4 and CAT5 questions assess this Expectation.

- Under Goal 2, the Academic Expectations numbered 2.24 through 2.38 are not assessed by very many KIRIS96 questions. They are also not assessed by any of the CTBS4 questions and only a few of the CAT5 questions. Some areas, such as *language appreciation* (2.27) and *second language* (2.28) tend not to be assessed at all.

- None of the CTBS4 and only very few of the CAT5 questions could be classified under Goal 5. Not as many KIRIS96 question as might be expected were classified under this goal, but there are at least some that are.

- Our method of classifying the questions for Goal 5, however, may not accurately reflect the abilities students are required to use. Initially we classified very few of the questions on any of the tests as focusing *primarily* on assessing critical thinking, creativity, and problem solving as these are defined in the Kentucky curriculum framework for the different grade levels. However, we used the definitions of critical thinking and problem solving in the curriculum framework at a particular grade level. Thus, if an eleventh-grade question did not require the student to use problem solving or critical thinking skills and processes in the same way they are defined for that grade level, it was not classified in one of these Academic Expectation categories. This resulted in not classifying a question in Goal 5, even though the question seemed to require a student to use a lower grade level expectations for problem solving or critical thinking. Also, the general scoring rubrics provided to us were not specific enough for us to decide whether students were actually going to be evaluated on the problem solving and critical thinking abilities defined by the Academic Expectations. Further, the definitions of Academic Expectations in the curriculum framework sometimes imply that a performance question is required rather than a paper-and-pencil question . Thus, one could argue that under other definitions or broader interpretations, many of the KIRIS96 questions do require students to use the abilities defined by Goal 5's Academic Expectations. Because of these considerations, we reexamined the questions and used a somewhat less stringent interpretation to classify the questions. The results we report reflect this less stringent approach Our findings regarding Goal 5, therefore, may not be correct.

- Goal 6 is another area that tends to have few KIRIS96 questions assessing it. None of the

CTBS4 and CAT5 questions appear to assess this Goal.

- On the whole, the complexity and intellectual richness of the KIRIS96 questions stand in contrast to the more narrowly focused multiple-choice questions of the standardized tests. These latter test questions, however, often do assess students' ability to use higher-order thinking skills such as making comparisons, drawing conclusions, and making inferences. Nevertheless, they do not require students to use complex cognitive integration in the way that the KIRIS96 questions do.

- There are some areas that the standardized tests evaluate separately and which the KIRIS96 does not evaluated separately. These areas might be called the "traditional basic skills" areas. As used in this report, the term *traditional basic skills* means the areas of simple reading comprehension, writing, spelling, mathematics computation, basic knowledge of mathematics concepts, language mechanics, and use of specific language expressions. The CTBS4 and the CAT5 assess, albeit in more simplified ways than the KIRIS96, and provide separate scores for the areas of *spelling*, *mathematics computation*, *basic knowledge of mathematics concepts*, *language mechanics*, and *use of specific language expressions*.

- The KIRIS96 Reading assessment does assess reading comprehension, and the KIRIS96 writing on-demand assessment and the writing portfolio do provide measures of the basic writing skills. However, even though language usage and spelling influence the scores students receive in writing, there are no separate scores for these areas. Even if a student's writing on the KIRIS96 were to be scored for spelling or language mechanics, such a score would be based only on the particular words and language a student chose to use when writing the response. A specially constructed test of spelling and language mechanics has the advantage of systematically including questions from the full range of relevant spelling patterns and language forms for a grade. This gives a better indication of a student's abilities in these areas than assessing only what a student chose to use in a written response. (This statement should be interpreted as a statement of a general principle of assessment and not as an endorsement of the CAT5 or the CTBS4 as good measures of spelling or language usage.)

Some Possible Implications of the Study

- If traditional basic skills as defined here are considered along with the Kentucky Academic Expectations, then we conclude that the KIRIS96 open-response questions and the standardized tests complement one another. Using the assessment approaches of both procedures would provide information on the attainment of the traditional curriculum's basic skills as well as on the attainment of the Academic Expectations.

- The *TerraNova*, a new CTB/McGraw-Hill product, includes a set of Supplemental Tests, Called the *Plus*. These tests assess word analysis, vocabulary, language mechanics, spelling, and mathematics (computation). Schools who value these traditional curriculum basic skills may wish to consider this product as a supplement to the KIRIS tests. (This is not an endorsement of the *Plus*. Other test publishers may have products that are equal or better than the *Plus*. We have not evaluated other products.) See the Kentucky Institute for Educational Research's (KIER) *Guide to Tests in Kentucky* (1997) for a more complete comparison of this assessment option with the CTBS5, CAT5, and KIRIS.

- KIRIS96 does not assess Kentucky's Goal 5, Goal 6, and Academic Expectations numbered 2.23 and 2.38 as thoroughly as might be desired. The CTBS4 does not appear to assess these goals at all. The CAT5 does so minimally, but less well than the KIRIS96. If these parts of the curriculum framework are judged to be important, some additional focus on assessing them seems to be warranted.

- Except as noted previously, the KIRIS96 samples the Kentucky Expectations reasonably well at the school level. We note that at the individual student level, however, the KIRIS96 had too few questions administered to each student to thoroughly measure a student in each curriculum area. When there are too few questions per student, it raises a *validity issue*: "What are the limitations of interpreting an individual student's score as a measure of that student's mastery of the Kentucky Academic Expectations?" This is a matter for technical discussion and debate. However, our judgment is that there are too few questions to have much confidence in an individual student's score as a sound meas-

ure of that student's mastery of the Academic Expectations.

- The 1996-1997 KIRIS (KIRIS97) assessments contain a mixture of open-response and multiple choice questions. The total number of questions per student was thus increased. (See the KIER *Guide to Tests in Kentucky* (1997) for a concise description of the KIRIS97 and related standardized tests.) We suggest that new studies be conducted on the KIRIS97 questions to (1) evaluate the extent to which they match the Academic Expectations, (2) study the stability of the scores over a short interval of time, and (3) determine the extent to which an individual student's score is influenced by responding to different questions as described previously by the matrix sampling test design. Information from such studies would inform the policy decision of whether to use the KIRIS results in the future as individual student measures of achievement.

- We note that the KIRIS96 questions do not provide a thorough assessment of an individual student's strengths and weaknesses in relation to the different parts of the curriculum framework. Each student responds to too few questions in relation to the curriculum framework. As a result, a student's score is likely to be too highly influenced by the specific topics and content that happen to be the focus of those few specific questions. Whether the KIRIS97, being longer, is valid for measuring a student's strengths and weaknesses regarding different parts of the curriculum, needs to be informed by research. This matter of too few questions for individual assessment is a *validity issue* that should be a matter for research so that data can be used to inform the discussion and policy decisions.

- Unlike a standardized test, the KIRIS assessment does not measure an individual student's growth from year-to-year along a psychometrically constructed achievement continuum. The current assessment design would have to be changed if such individual student assessment is required. An issue for discussion and technical research is the validity and reliability of measuring an individual student's growth using the KIRIS and a standardized test.

- The CTBS4 and the CAT5 were not constructed specifically to meet the Kentucky curriculum framework so it may be unreasonable to expect them to match the framework well. Further, they use traditional content rather than the new approaches recommended by professional associations as a framework for developing questions. The KIRIS96 and the new CTBS Fifth Edition, however, use the newer approaches to curriculum content as a basis of test development. Since the Fifth Edition of the *Comprehensive Tests of Basic Skills*, Survey Edition, (CTBS5) (Reading/Language Arts and Mathematics subtests only) were required to be administered in Kentucky in 1997, it is recommended that a study be undertaken to evaluate the extent to which the CTBS5 and the KIRIS97 match the Academic Expectations.

Limitations of Our Study

Our results and conclusions have certain limitations which we discuss below.

- The assessment questions were classified using holistic judgments. We did not develop an objective way of classifying questions. Thus, our classification results may be subject to inconsistencies and errors which our best efforts may not have controlled well enough.

- Although questions from different tests are classified into the same Academic Expectation category, one should not infer that the questions or tests are identical. Each assessment instruments has a unique way of posing a question or setting a question. These different ways result in different kinds of thinking being used by the students. Our results give only a rough guideline as to the content areas covered by the different assessments.

- We did not perform an analysis to match the KIRIS96 open-response questions with the test specifications of the CTBS4 or the CAT5. Thus, this study presents only one side of the picture. However, such a matching study may be of little use, since the main issue is whether the tests match the Kentucky curriculum framework.

- Our classification of test questions was not based on a deep psychological study of the cognitive skills and mental processes that students use to respond to the questions. Such a study might have revealed a classification that is quite different from ours.

- As was discussed previously, our interpretation of the Academic Expectations under Goal 5 may not be correct.

- Sometimes we found that either the framework definitions were not clear or a KIRIS96 open-response question did not appear to be crafted with the Academic Expectation specifically in mind. (None of the documentation of the KIRIS96 question development process clearly indicated the path that was followed in deriving the questions from and assuring that they matched the Academic Expectations.) Thus, we often had to make assumptions about the main intent of the questions in relation to the Academic Expectations. To the extent that our assumptions are in error, to that extent are the questions misclassified.

Table 1. Percentage of fourth grade assessment questions classified into each Academic Expectation category

Grade 4	KIRIS96	CTBS4	CAT5		KIRIS96	CTBS4	CAT5		KIRIS96	CTBS4	CAT5
ACADEMIC EXPECTATIONS				ACADEMIC EXPECTATIONS				ACADEMIC EXPECTATIONS			
GOAL 1				2.8 Mathematical procedures	5%	13%	12%	2.30 Consumerism	2%		1%
1.1 Accessing information	1%	10%	13%	2.9 Space & dimensionally	5%	1%	2%	2.31 Physical wellness	5%		
1.2 Reading*	17%	17%	10%	2.10 Measurement	1%		1%	2.32 Mental wellness			
1.3 Observing			3%	2.11 Change	1%			2.33 Health systems			
1.4 Listening				2.12 Mathematical structure				2.34 Psychomotor skills	1%		
1.5-1.9 Math procedures	2%			2.13 Data	2%	2%	2%	2.35 Physical activities	1%		
1.10 Classifying			7%	2.14 Democratic principles	4%			2.36 Career path	3%		
1.11 Writing	100%			2.15 Political systems	1%	1%		2.37 Transition skills	1%		
1.12 Speaking				2.16 Social systems	10%		3%	2.38 Post secondary search	1%		
1.13 Visual arts				2.17 Cultural diversity	1%						
1.14 Music				2.18 Economic systems	5%			GOAL 5			
1.15 Movement				2.19 Geography	4%	2%	5%	5.1 Critical thinking***	5%		2%
1.16 Technology				2.20 Historical perspective	2%		3%	5.2 Creative thinking***	5%		
				2.22 Production	4%		2%	5.3 Conceptualizing***	1%		1%
GOAL 2				2.23 Analysis of form	7%			5.4 Decision making***			
2.1 Scientific Skills**	12%	9%	10%	2.24 Aesthetics	4%			5.5 Problem solving***	4%		
2.2 Patterns	4%			2.25 Cultural heritage							
2.3 Systems and interactions	6%	1%		2.26 Cultural diversity	2%			GOAL 6			
2.4 Models and scales	2%			2.27 Language				6.1 Multiple perspectives	1%		
2.5 Constancy	3%			2.28 Second language				6.2 Acquire knowledge	1%		
2.6 Evolutionary change	1%			2.29 Family life	4%		2%	6.3 Expanding knowledge	1%		
2.7 Number concepts**	5%	11%	10%								
								WERE NOT CLASSIFIED	3%	28% ****	30% ****
								Total number of distinct questions	164	408	200

* The CTBS4 and CAT5 vocabulary questions were classified into the reading category.

** Many of the science and mathematics concept questions fit the scientific skills and number concepts Academic Expectations only approximately. The questions tend to be too narrow to fit the complexity implied by the Academic Expectations. However, they seem to assess related, but lower-order thinking skills. Many science questions on CTBS and CAT seem to assess scientific concepts rather than scientific skills.

*** The way we judged the KIRIS96 questions in this area may not be correct. See the text for an explanation.

**** The CTBS4 and CAT5 contain language mechanics, language expression, and spelling questions. These were not classified into any Academic Expectations categories.

Table 2. Percentage of eighth grade assessment questions classified into each Academic Expectation category

Grade 8 ACADEMIC EXPECTATIONS	KIRIS96	CTBS4	CAT5	ACADEMIC EXPECTATIONS	KIRIS96	CTBS4	CAT5	ACADEMIC EXPECTATIONS	KIRIS96	CTBS4	CAT5
GOAL 1				2.8 Mathematical procedures	8%	11%	10%	2.30 Consumerism	3%		2%
1.1 Accessing information		8%	10%	2.9 Space & dimensionally	6%	1%	2%	2.31 Physical wellness	4%		
1.2 Reading*	18%	15%	13%	2.10 Measurement	4%			2.32 Mental wellness	2%		
1.3 Observing				2.11 Change				2.33 Health systems	3%		
1.4 Listening				2.12 Mathematical structure	4%			2.34 Psychomotor skills	1%		
1.5-1.9 Math procedures	6%			2.13 Data	5%	2%		2.35 Physical activities	2%		
1.10 Classifying			11%	2.14 Democratic principles	4%			2.36 Career path	3%		
1.11 Writing	100%			2.15 Political systems	4%		1%	2.37 Transition skills	2%		
1.12 Speaking				2.16 Social systems	7%	2%	2%	2.38 Post secondary search	1%		
1.13 Visual arts				2.17 Cultural diversity	5%						
1.14 Music				2.18 Economic systems	5%		2%	GOAL 5			
1.15 Movement				2.19 Geography	4%	2%	3%	5.1 Critical thinking****	7%		
1.16 Technology				2.20 Historical perspective	1%	3%	4%	5.2 Creative thinking***	4%		
				2.22 Production	4%			5.3 Conceptualizing***	1%		
GOAL 2				2.23 Analysis of form	6%		2%	5.4 Decision making***	2%		
2.1 Scientific Skills**	1%	10%	10%	2.24 Aesthetics	4%		2%	5.5 Problem solving***	9%		
2.2 Patterns	8%			2.25 Cultural heritage	4%						
2.3 Systems and interactions	2%			2.26 Cultural diversity	1%			GOAL 6			
2.4 Models and scales	2%			2.27 Language				6.1 Multiple perspectives	1%		
2.5 Constancy	1%			2.28 Second language				6.2 Acquire knowledge	1%		
2.6 Evolutionary change	7%			2.29 Family life	2%			6.3 Expanding knowledge	1%		
2.7 Number concepts**	7%	11%	10%								
								WERE NOT CLASSIFIED		28%****	30%****
								Total number of distinct questions	168	408	200

* The CTBS4 and CAT5 vocabulary questions were classified into the reading category.

** Many of the science and mathematics concept questions fit the scientific skills and number concepts Academic Expectations only approximately. The questions tend to be too narrow to fit the complexity implied by the Academic Expectations. However, they seem to assess related, but lower-order thinking skills. Many science questions on CTBS and CAT seem to assess scientific concepts rather than scientific skills.

*** The way we judged the KIRIS96 questions in this area may not be correct. See the text for an explanation.

**** The CTBS4 and CAT5 contain language mechanics, language expression, and spelling questions. These were not classified into any Academic Expectations categories.

Table 3. Percentage of eleventh grade assessment questions classified into each Academic Expectation category

Grade 11	KIRIS96	CTBS4	CAT5		KIRIS96	CTBS4	CAT5		KIRIS96	CTBS4	CAT5
ACADEMIC EXPECTATIONS											
GOAL 1				2.8 Mathematical procedures	9%	11%	10%	2.30 Consumerism			1%
1.1 Accessing information		7%	10%	2.9 Space & dimensionally	4%	1%	2%	2.31 Physical wellness	2%		
1.2 Reading*	21%	32%	23%	2.10 Measurement	2%			2.32 Mental wellness	1%		
1.3 Observing				2.11 Change				2.33 Health systems	1%		
1.4 Listening				2.12 Mathematical structure				2.34 Psychomotor skills			
1.5-1.9 Math procedures				2.13 Data	6%	1%		2.35 Physical activities			
1.10 Classifying			4%	2.14 Democratic principles	1%	1%		2.36 Career path	2%		
1.11 Writing	100%			2.15 Political systems	4%		1%	2.37 Transition skills	1%		
1.12 Speaking				2.16 Social systems	2%	1%	5%	2.38 Post secondary search	2%		
1.13 Visual arts				2.17 Cultural diversity	1%	2%					
1.14 Music				2.18 Economic systems	8%		2%	GOAL 5			
1.15 Movement				2.19 Geography	4%	4%	2%	5.1 Critical thinking***	10%		
1.16 Technology				2.20 Historical perspective	5%	2%	1%	5.2 Creative thinking***	7%		2%
				2.22 Production	4%			5.3 Conceptualizing***	1%		1%
GOAL 2				2.23 Analysis of form	10%			5.4 Decision making***	2%		
2.1 Scientific Skills**	11%	9%	10%	2.24 Aesthetics	2%		1%	5.5 Problem solving***	2%		
2.2 Patterns	4%		1%	2.25 Cultural heritage	1%						
2.3 Systems and interactions	10%		1%	2.26 Cultural diversity	1%			GOAL 6			
2.4 Models and scales	1%			2.27 Language				6.1 Multiple perspectives	3%		
2.5 Constancy	1%		2%	2.28 Second language				6.2 Acquire knowledge	1%		
2.6 Evolutionary change	2%			2.29 Family life	1%			6.3 Expanding knowledge			
2.7 Number concepts**	2%	20%	15%								
								WERE NOT CLASSIFIED		28% ****	30% ****
								Total number of distinct questions	168	408	200

* The CTBS4 and CAT5 vocabulary questions were classified into the reading category.

** Many of the science and mathematics concept questions fit the scientific skills and number concepts Academic Expectations only approximately. The questions tend to be too narrow to fit the complexity implied by the Academic Expectations. However, they seem to assess related, but lower-order thinking skills. Many science questions on CTBS4 and CAT5 seem to assess scientific concepts rather than scientific skills.

*** The way we judged the KIRIS96 questions in this area may not be correct. See the text for an explanation.

**** The CTBS4 and CAT5 contain language mechanics, language expression, and spelling questions. These were not classified into any Academic Expectations categories.